

LA REALIDAD AUMENTADA: UNA TECNOLOGÍA EN ESPERA DE USUARIOS

Biol. Lizbeth Heras Lara

*Técnico académico, Departamento de Visualización, DCI, DGSCA,
UNAM*

Biol. José Luis Villarreal Benítez

Técnico académico, SERUNAM, DGSCA, UNAM

LA REALIDAD AUMENTADA: UNA TECNOLOGÍA EN ESPERA DE USUARIOS

Resumen:

Este trabajo se aborda el tema de la Realidad Aumentada, que es una estrategia tecnológica poco conocida y usada, debido a que para su uso es necesario integrar varias tecnologías, lo cual se ha hecho a través de dispositivos especializados (lentes de realidad aumentada, que incluyen como base un sistema de captura de video y uno de proyección, los cuales son integrados por software en una computadora equipada con hardware gráfico).

Además, abordamos el tema de las tecnologías para la representación tridimensional de datos e ideas y la interacción humano-máquina más sofisticada; sus procesos para entender cómo funciona; los avances obtenidos y los que queda pendiente; y se ofrecen ideas para iniciar proyectos como desarrolladores de aplicaciones o de contenidos.

Palabras clave: Realidad aumentada, interacción, hardware, gráficos

AUGMENTED REALITY: A TECHNOLOGY AWAITING USERS

Abstrac:

This work approaches subject of Reality Increased, that is strategy technological little well-known and used, because for his use it is necessary to integrate several technologies, which has done through devices specialized (lenses of reality increased, that includes as it bases a system of capture of video and one of projection, which are integrated by software in a computer equipped with graphical hardware).

In this text we approached the subject of the technologies for the three-dimensional representation of data and ideas and the interaction more sophisticated human-machine; its processes to understand how it works; the obtained advances and those that slope is left; and ideas are offered to initiate projects like content or application developer.

Key words: Increased reality, interaction, hardware, graphs

Introducción

La síntesis de imágenes por computadora ha sido una herramienta que científicos, ingenieros y humanistas han utilizado para representar sus datos e ideas visualmente. También ha sido empleada como una forma de interactuar con las computadoras a través de metáforas y de representaciones de mundos o espacios concebidos por ellos, aún no construidos, imposibles de construir en el mundo real, pero que al existir en el mundo virtual adquiere un significado o valor especial.

La tendencia a representar los resultados de los cálculos de las computadoras con imágenes, muchas de ellas tridimensionales, así como la facilidad para interactuar con las máquinas a través de metáforas gráficas, ha permitido el desarrollo de software y hardware muy sofisticados que permiten la síntesis de imágenes tridimensionales con extremo detalle visual, y a velocidades tales que engañan al ojo humano, generando sensaciones de realismo.

El realismo alcanzado por los sistemas visuales (videojuegos, salas de realidad virtual, teatro-domos, simuladores gráficos, entre otros) se basa en el detalle gráfico o foto-realismo, en la simulación de las reglas de movimiento del mundo real (principalmente la física de fuerzas, gravedad, fricción), y en la respuesta de los elementos en el mundo virtual ante los estímulos o comandos del usuario (la interactividad).

Estos ambientes gráficos van desde mundos virtuales, que nada tienen que ver con nuestros referentes reales y en los que las geometrías y las reglas que regulan a los objetos responden a conceptos abstractos, hasta reconstrucciones o réplicas virtuales de espacios que existen o existieron y sobre los cuales un usuario reproduce las mismas sensaciones que le generarían en el mundo real (ahora inaccesible por razones geográficas o del tiempo).

Sin embargo, las exigencias de los usuarios y las particularidades de cada problema; así como las limitaciones de lo real (zonas arqueológicas en ruinas, artefactos inconclusos, etc.) y lo virtual (generación limitada de estímulos para todos los sentidos), han consolidado estrategias que toman e integran lo mejor de ambos mundos. Las tecnologías para mezclar realidades como la realidad mixta (integrar en un espacio coherente para el usuario, a espacios sintetizados por computadora y espacios físicos donde se sitúa en ese instante el usuario) responde a este reto y aprovecha las tecnologías de tarjetas gráficas, despliegues de alta resolución y monitores miniatura, tecnologías móviles y de localización, cámaras de video digital pequeñas, tecnologías Web e inalámbricas; así como software de visualización desarrollado en las últimas décadas.

La integración de tantas tecnologías para resolver un reto tecnológico parece un reto inalcanzable. No obstante, las estrategias de convergencia tecnológica y estandarización están alcanzando una madurez tal que su explotación es cada día más sencilla. Por otro lado, el desarrollo de aplicaciones y contenidos en Realidad Aumentada y Virtual se basa en la integración de herramientas desarrolladas en la visualización científica en los últimos treinta años y en herramientas específicas de la realidad mixta en los últimos quince años.

En este trabajo nos concentramos en el área de realidad aumentada. Esta es una tecnología muy poco conocida usada, debido a la necesidad de integrar varias tecnologías, a través de dispositivos especializados (lentes de realidad aumentada, los cuales incluyen como base un sistema de captura de video y uno de proyección, los cuales son integrados por software en una computadora equipada con hardware gráfico). Sin embargo, es posible aprovechar muchos dispositivos ya disponibles en un sector amplio y con tendencia de mayor penetración (PDA, Celulares, TabletPC, principalmente) que ya integran los diversos componentes requeridos y pueden ser adaptados para montarles contenidos con Realidad Aumentada. Esta estrategia académica permitirá estar preparados para la distribución de contenidos, servicios y aplicaciones de realidad aumentada o complementada.

¿Cómo funciona?

La integración de objetos y mundos reales y virtuales, a veces agregados, combinados o fusionados o intercambiados, es el área de la creación y manejo de mundos integrados o realidad mezclada. Esta área de la visualización se basa en una estrategia de visualización e interactividad que hace uso de muchas tecnologías y de diferentes áreas de visualización científica; dando lugar a un espectro de modalidades que se mueven alrededor de dos extremos: si el entorno del espacio (el ambiente circundante) es predominantemente virtual y se le agregan objetos virtuales y reales, se habla de realidad virtual; mientras que si el entorno dominante es real y se le agregan objetos virtuales, se habla de realidad aumentada.

La realidad aumentada es una tecnología que integra señales captadas del mundo real (típicamente video y audio) con señales generadas por computadores (objetos gráficos tridimensionales); las hace corresponder para construir nuevos mundos coherentes, complementados y enriquecidos – hace coexistir objetos del mundo real y objetos del mundo virtual en el ciberespacio-.

Esta tecnología aprovecha las tecnologías derivadas de la visualización para construir aplicaciones y contenidos con las cualidades que estas áreas han madurado en las últimas décadas. Del procesamiento de imágenes toma la cualidad de resaltar aspectos en las imágenes captadas por la cámara de video, estos rasgos son analizados por procesos de visión para extraer propiedades geométricas del entorno y los objetos (posición tridimensional, patrones fiduciaros para el reconocimiento y ubicación de objetos susceptibles a sustitución, etc.). De los gráficos por computadora toma la síntesis de objetos tridimensionales y sus transformaciones, mientras que gracias a la teoría de interfaces gráficas ha sido posible la construcción de nuevas metáforas dentro de estos mundos mixtos.

Un sistema de realidad aumentada general inicia con el registro de las señales del mundo real (video y audio, aunque se continúan evaluando subsistemas para la síntesis de señales para los otros sentidos). Estas señales son procesadas por un sistema de realce de orillas de objetos para preparar la imagen para la segmentación o extracción de objetos y el reconociendo de patrones y marcas fiduciaras. Este proceso permite determinar en dónde hay que remplazar un objeto real por uno virtual, cuál objeto virtual debe colocarse sobre el espacio real (el espacio de video) y en qué posición y perspectiva.

Para agilizar el proceso y permitir la interactividad, la cual requiere de gráficos en tiempo real, es conveniente que la correspondencia entre patrones, marcas fiduciaras, rasgos geométricos del entorno, y la posición tridimensional y la perspectiva de dibujo de los objetos virtuales, sea preparada con anticipación. Esto es: que se cree una base de datos y se entrene al sistema para evitar muchos cálculos en tiempo de ejecución. Algunos de estos cálculos pueden ser preparados y dicha inteligencia artificial ser parte del sistema. Sin embargo, la síntesis de imágenes en su posición y perspectiva correctas, que permita lograr una correspondencia geométrica entre los mundos virtual y real (que el nuevo espacio se coherente para usuario), debe realizarse en tiempo real y de forma interactiva. Esta alineación de ambos mundos se logra extrayendo información tridimensional de las imágenes de video (en dos dimensiones) a partir de marcas fiduciaras en el mundo real y de rasgos de perspectiva que pueden ser extraídos del entorno real (los contornos de muros, geometrías simples y conocidas o medidas previo a la generación del contenido, entre otras técnicas). Esta característica restringe las aplicaciones de realidad aumentada en mundos virtuales específicos a mundos reales; de otra forma, la demanda de cálculo exige computadoras poderosas, de tal suerte que estos casos deben esperar algunos años.

Cuando se conocen las propiedades del dibujo (se establecen la serie de transformaciones que han de aplicarse al objeto virtual; así como los parámetros de la cámara virtual y la iluminación), se crean o sintetizan y se pasan al proceso de composición de la señal de video de salida; la cual fusiona siguiendo reglas (de oclusión, por ejemplo) la señal de video original con la señal de la escena virtual. Esta nueva señal, mezcla de ambos mundos, es transferida a los monitores o proyectores.

Por último, esta señal que contiene una reconstrucción visual (sonora y cualquier otra señal registrada o sintetizada) de la escena objetivo de la aplicación, es dirigida al sistema visual humano. Si el sistema genera una perspectiva única para ambos ojos, el usuario verá una imagen bidimensional (mono); mientras que para ver en tercera dimensión (estereografía) es necesario generar un par de imágenes, cada una con la perspectiva correspondiente a cada ojo y solo dejar ver a cada ojo su imagen correspondiente. Lo que ve el usuario es una interpretación producto de un proceso neuro-psicológico.

La concatenación de estos procesos resulta en un sistema con las siguientes características, las cuales definen la realidad aumentada:

- Combina objetos reales y virtuales en nuevos ambientes integrados.
- Las señales y su reconstrucción se ejecutan en tiempo real.
- Las aplicaciones son interactivas.
- Los objetos reales y virtuales son registrados y alineados geoméricamente entre ellos y dentro del espacio, para darles coherencia espacial.

El estado actual

Los inicios de la realidad aumentada se remontan a 1960, cuando Sutherland usó un dispositivo de despliegue de imágenes tridimensionales de tipo casco, para visualizar gráficos tridimensionales. Sin embargo, esta tecnología se ha desarrollado en los últimos quince años; ha madurado en hardware, software, aplicaciones y contenidos. Actualmente, muchos celulares, consolas de videojuegos, PDA y TabletPC ya cuentan con algunos de los dispositivos necesarios para implementar realidad aumentada.

Los dispositivos especializados, los cuales integran el hardware necesario, más usados son los lentes con monitores montados sobre un armazón tipo gafas, con una cámara tipo webcam colocada sobre el soporte entre los lentes. Estos lentes cuentan con controladores que reconocen diferentes herramientas de software, tales como MagicBook y ARToolkit. Sin embargo, muchos dispositivos móviles permiten montar aplicaciones de realidad aumentada; por ejemplo, las PDA cuentan con cámara de video y pantallas orientadas ortogonalmente y alineadas, apuntando uno hacia delante y el otro hacia atrás; lo cual permite configurar una herramienta de realidad aumentada. Las TabletPC también pueden ser configuradas de forma similar; algunas cuentan con cámara y en caso contrario, basta con fijar una webcam en la orientación adecuada.

La mayoría de estos dispositivos cuentan con intercomunicadores basados en radiofrecuencia (Wi-Fi, BlueTooth, Wi-Max), por lo que pueden procesar la intensidad de las señales recibidas por los diferentes puntos de acceso (switches para acceso a Internet y otros dispositivos) dentro de una red inalámbrica (por ejemplo la RIU dentro de la UNAM) para determinar su ubicación dentro del entorno real; si se les proporciona la localización geográfica (o espacial local) de los diferentes puntos de acceso (wireless access points).

Aunque es posible usar muchos dispositivos, que ya han penetrado el mercado y sus costos son bajos, considerando su multifuncionalidad, los procesadores aún son lentos para las tareas que se requieren en realidad aumentada y muchos de ellos carecen de hardware para aceleración gráfica (aunque la tecnología ya está lista y disponible, aún es cara; por ejemplo, algunos teléfonos celulares cuentan con el coprocesador para gráficos 3D - ATI Imageon 2300), por lo que las aplicaciones y los contenidos dependen de una preparación exhaustiva de los entornos reales. Por ejemplo, con marcas fiduciarias y patrones simples y muy diferenciados entre ellos. Las aplicaciones independientes de los entornos reales o con entornos muy diversos, tendrán que esperar algunos años.

Estos dispositivos genéricos y los especializados en realidad aumentada, están listos y hay muchas aplicaciones evaluadas, pero sólo dan soporte para video, audio y gráficos tridimensionales. Poco se ha avanzado en otros sentidos (táctil y equilibrio, principalmente), fuera de los trabajos académicos en laboratorios dedicados a la realidad mixta. Sin embargo, es posible recurrir a la reserva de algoritmos, aplicaciones y experiencia de la visualización científica, de donde se ha nutrido la realidad aumentada, y enriquecer nuestros contenidos.

La problemática actual para desarrolladores de aplicaciones y contenidos es:

- a) la construcción de software modular listo para diferentes dispositivos o hardware (deploy - componentes estándares listos para su instalación, prueba y configuración en diferentes arquitecturas de hardware),
- b) el desarrollo de contenidos estándares independientes de las aplicaciones y con riqueza de representación.

Así mismo, ambos creadores deben recurrir a algoritmos para soluciones alternativas o novedosas que permitan la implementación de los nuevos retos. La realidad aumentada se nutrió de los avances de la visualización como disciplina; es tiempo de revisar nuevamente sus avances e incorporarlos a la disciplina de la realidad mixta y en particular de la realidad aumentada, estas áreas siguen siendo parte de la visualización científica.

En este sentido, es importante notar que actualmente la visualización no sólo se ocupa del análisis y síntesis de objetos visuales percibidos por el sistema visual humano (aunque fueron las señales ópticas las primeras y quizá las dominantes hasta el momento), también trata con las señales que perciben los demás sentidos, tanto por su estrecha relación con el sistema visual, como por las similitudes en los procesos físicos y neurofisiológicos. Por otro lado, los avances en el hardware especializado en gráficos han permitido la implementación de técnicas y algoritmos en tiempo real. Este conocimiento desarrollado para tareas específicas de la visualización, principalmente en procesamiento de imágenes, visión por computadora y render, está siendo retomado por los desarrolladores de aplicaciones y contenidos en Realidad Aumentada.

Más específicamente, algunos problemas que persisten y siguen siendo retos para el desarrollo de la tecnología o los principales puntos de falla de los contenidos y ejemplos, son los siguientes:

1. Problemas de paralaje por el desfase de la cámara con respecto a los ojos.
2. El seguimiento de los ojos para ampliar el campo de visión y aprovecharlo como interfaz de interacción humano-máquina.
3. La oclusión o eliminación de los objetos reales que se sustituyen con los virtuales, o viceversa.
4. La complejidad de ambientes exteriores o la extensión de una aplicación o contenidos a ambientes no preparados.
5. El retraso del cálculo de transformaciones geométricas con respecto al despliegue de video.
6. Los cálculos de localización.

En seguida se presentan algunos detalles de dicha problemática.

La colocación de una cámara de video para capturar una escena del mundo real, generalmente no corresponde con la vista del los ojos, y produce un error de paralaje. Este problema está más marcado en la adaptación de dispositivos no especializados en realidad aumentada; las compañías de celulares y PDAs no están considerando la realidad aumentada en la construcción de sus dispositivos, pero ésto cambiará cuando se produzcan contenidos y aplicaciones interesantes para dichos dispositivos.

Los dispositivos actuales, incluso los especializados o diseñados específicamente para realidad aumentada, tienen un campo de visión limitado (alrededor de 30 grados). Para ampliar el campo de visión no sólo se requieren monitores más grandes, también es importante seguir electrónicamente la posición de los ojos (conocer con precisión hacia dónde está mirando el usuario). Los dispositivos que cubren esta función, ya son accesibles y baratos, inclusive se han alcanzado buenos rendimientos con simples cámaras web (por ejemplo, ver el proyecto Dasher; <http://www.inference.phy.cam.ac.uk/dasher/>).

Otra solución al problema del campo de visión, es el uso de los lentes traslúcidos y el sistema de proyección sobre los lentes; esto es, ver a través del lente el mundo real y proyectar la señal virtual sobre el lente. Sin embargo, este sistema tiene poco contraste y problemas de oclusión. Aunque se consigue un campo de visión natural y no presenta problemas de paralaje.

El retraso por baja latencia es uno de los errores persistentes, para lo cual se requieren procesadores más potentes; el poder de cómputo en dispositivos móviles ya es alto y se sigue incrementando. Los trabajos recientes han mostrado que para este caso son útiles los algoritmos de predicción de movimiento, también es posible compensar movimientos pequeños del dispositivo con metamorfosis de imágenes (image warping), lo cual es más rápido que el render en tercera dimensión.

Por último, el reto de la realidad aumentada está en su uso sobre ambientes no preparados. No es práctico poner marcas en ambientes exteriores y es necesario buscar patrones fácilmente reconocibles, como son las siluetas en el paisaje y grandes estructuras simétricas. Para estos ambientes es importante incorporar tecnologías de localización espacial como los GPS (sistemas de posicionamiento global), combinados con bases de datos con sistemas de información geográficos.

Aplicaciones

El sistema visual humano, y los otros sentidos, perciben el mundo físico o real dentro de un contexto. Esta percepción es una reconstrucción, una interpretación; la realidad es multidimensional y multicontextual. La Realidad Aumentada permite desglosar los diferentes aspectos o modelos para que el sistema visual humano y los otros sentidos reciban los aspectos adecuados y en muchas ocasiones ocultos a los sentidos, así como modelos generados por modelos que simplifican la complejidad que la naturaleza multidimensional del mundo.

Las aplicaciones pertinentes de la realidad aumentada son aquellas que requieren la reformulación del mundo con información multidimensional, para presentar versiones reducidas y reestructuradas para revelar conocimiento. Las principales aplicaciones se han dado en campos muy diversos que tienen los aspectos arriba mencionados, en común. Los casos mejor documentados son en educación, el arte, entrenamiento industrial, entretenimiento, difusión de la ciencia y la tecnología, museos, presentación de productos, narraciones interactivas, y en la industria militar.

¿Qué podemos hacer? ¿Dónde iniciar?

La tradición de interfaces gráficas en RA ha sido la presentación de la mezcla de los mundos virtuales y reales sobre la pantalla y la interacción directa con el mundo real y el virtual. Aunque esta interacción es difícil, con la madurez en la tecnología seguramente se le incorporarán mejores interfaces gráficas. Sin embargo, un aspecto atractivo, que ha mantenido esta tendencia, es el uso de la RA como interfaz para la interacción humano-máquina, que es un buen punto de partida para la elección de esta tecnología como un medio para la producción de contenidos y aplicaciones.

La capacidad de la realidad aumentada para incorporar objetos reales y virtuales en un mismo escenario, permite el uso de artefactos reales para manipular objetos virtuales o acciones y comportamientos de los objetos virtuales de los procesos de la computadora; esto es, la capacidad de la RA para darle sentido a los rasgos de los objetos reales y transformarlos en virtuales o mixtos, permite usar esta tecnología como una plataforma para la interacción humano-máquina.

Si bien falta mucho trabajo para desarrollar esta tecnología, el intenso trabajo en el área durante los últimos quince años ha permitido montar aplicaciones útiles en educación y entretenimiento. Se ha realizado mucho esfuerzo para que los desarrollos de contenido usen las herramientas de manera muy sencilla, prácticamente integrando módulos. La realidad aumentada está lista, faltan usuarios curiosos que exploren la tecnología y fuercen a los desarrolladores para juntos llevarla al siguiente nivel.

Para participar en esta tarea, hay que iniciar por imaginar mundos que no pueden existir físicamente, ni en un espacio virtual; y construir historias como lo han hecho los pueblos a lo largo de los años: para su entretenimiento, para transferir sus tradiciones, enseñar a sus jóvenes y registrar sus hazañas.

Bibliografía

El software ARToolkit, el cual es una biblioteca para desarrollos rápidos de aplicaciones de realidad aumentada, es libre y está disponible en http://www.hitl.washington.edu/research/shared_space/.

Las principales aplicaciones se han dado en diferentes campos, se pueden apreciar algunas en:

http://studierstube.icg.tu-graz.ac.at/invisible_train/

<http://www.technologyreview.com/Infotech/18291/>

http://studierstube.icg.tu-graz.ac.at/handheld_ar/gizmondo.php

http://www.usc.edu/dept/architecture/mbs/thesis/anish/thesis_report.htm

http://studierstube.icg.tu-graz.ac.at/handheld_ar/virtuoso.php

<http://www.tinmith.net/>

http://www.ims.tuwien.ac.at/media/documents/publications/C3D_Video_Description-Final.pdf

<http://www.technologyreview.com/Biztech/17807>

<http://campar.in.tum.de/Chair/TeachingSs04SeminarAR>

<http://ieeexplore.ieee.org/iel5/8784/27815/01240727.pdf>

* Se anexa otro documento con las imágenes que ilustran el tema.